General solution for three-dimensional surface structures using direct methods

L. D. Marks
Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208
(Received 31 July 1998)

A general method for solving surface structures in three dimensions using surface x-ray-diffraction data coupled with direct methods is outlined. The method exploits the existence of a support constraint normal to the surface, and couples the concepts of projections, operators, and sets used in the image reconstruction literature with statistical operators used in direct methods. The approach presumes nothing beyond the fact that the scattering comes from atoms, and is a true model-independent approach. [S0163-1829(99)06324-9]

I. INTRODUCTION

For the last 20 or more years the standard approach when trying to determine a surface structure has been to compare experimental data against one or more models. This approach has a fundamental flaw; almost any technique [e.g., low-energy electron diffraction (LEED), reflection high-energy electron diffraction (RHEED), x-ray diffraction, ion scattering] can be refined to find a local minimum, and unless one of the models is an approximation to the true structure there is no reason for the comparison to be particularly relevant. In some cases scanning tunneling microscopy (STM) or high resolution electron microscopy (HREM) images may be available, both of which can yield important constraints on the symmetry and in some cases the atomic positions. In certain cases Patterson maps from x-ray or transmission electron diffraction experiments can be used, but with more complicated structures these can become hard if not impossible to decode.

It follows that the central issue in determining a surface structure is not in fact refining against some initial model, but instead finding these initial models. While enough is sometimes known about the general surface chemistry of a material to make knowledgeable guesses, guessing is not science. What is needed is a global search mechanism that can determine the set of feasible models for subsequent refinement. This is true independent of whether the experimental data is LEED I-V curves, RHEED rocking curves, or surface x-ray-diffraction data, the latter being the focus of this paper.

This problem is not unique to surface structure determination, but also occurs in bulk three-dimensional crystallography. An x-ray-diffraction experiment only measures the intensity in reciprocal space, from which the moduli of the structure factors can be obtained. Missing are the phases of the structure factors, and if these are known even relatively approximately a viable estimate of the charge density can be generated. In turn, peaks in the charge density can be associated with atomic sites and feasible models constructed.

To solve the phase problem, what are called ''direct methods'' have been developed over the last 30 years.[1–7] Additional information is available; the diffraction from x rays occurs from atoms, and the controlling equations are simple and cleanly defined. By building in the fact that the scattering comes from atoms statistical relationships linking the phases (primarily of the stronger reflections) are known to exist. Coupling these statistical relationships with experimental data permits construction of feasible models for the structure (often only one) for subsequent refinement. In essence, direct methods perform a global search for the initial models, eliminating guesswork.

While direct methods have been used for many years for three-dimensional structure determination, it is only very recently that they have been successfully employed for surface-structure determination,[8–17] and to date only with two-dimensional data. While there are problems when surface reflections overlap bulk reflections, in general a large set of measurements is available in two dimensions enabling solutions to be obtained. In three dimensions a much smaller fraction of all possible measurements to a reasonable resolution, e.g., 1 Å, is typically available. Perpendicular to the surface for a given rel-rod (different l values for given h and k in reciprocal space) the sampling of measurements along l is generally relatively small, corresponding to an effective c lattice parameter of 40–60 Å, and limited to resolutions of about 1.5 Å at best. Even excluding the case when (h,k,l) overlaps with a bulk reflection, typically only a small number of rods are known, less than 30% of all possible in some cases. With such incomplete data performing direct methods in three dimensions for surfaces would appear to be much harder, if not impossible.

The intention of this note is to demonstrate that this is not in fact the case, and three-dimensional surfaces can be solved using direct methods relatively straightforwardly. The main reason for this is the relatively fine sampling normal to the surface. Even though the effective unit cell parameter normal to the surface is large, 40–60 Å, substantial relaxations of atomic positions only take place in a relatively small region near the surface. Hence the charge density normal to the surface must be zero except in some small region of perhaps 10–20 Å total thickness. This is equivalent to what is called a ''support constraint'' in the image reconstruction literature,[18–21] and has a number of well established mathematical properties. Adding in this support constraint as additional information compensates for the small number of measurements in many cases.

The structure of this paper is as follows. In Sec. II, a brief outline of the basic approach is presented, a hybrid of classical direct methods and what is called the feasible set approach[22–25] from the image reconstruction literature which uses set theory, mathematical projections, and operators. (A

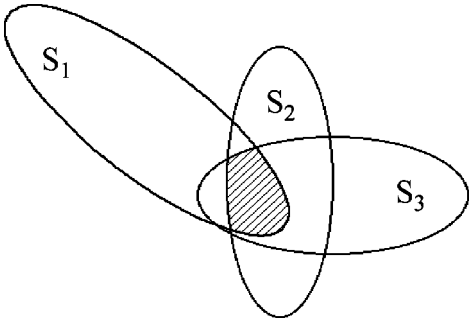

FIG. 1. Feasible set solution as the intersection of three convex sets, for instance the measured data, support and functional section as discussed in the text.

general discussion of a feasible set approach to direct methods is described elsewhere).[26] The theory is developed first for the case when the only rods which do not overlap with bulk reflections are available. This is then extended to the case when bulk data is available as well. Section III discusses an important exception to the general case, what we refer to as Babinet solutions, as well as a method of projecting onto them. This is then followed by numerical examples, concluding with a discussion.

II. THEORY

We have experimental measurements of the magnitudes of the structure factors, $|F(\mathbf{k})|$ for different three-dimensional reciprocal-lattice values $\mathbf{k}$. The aim is to find a good approximation to the phases. With this achieved, we can Fourier transform $F(\mathbf{k})$ to obtain an approximation of the surface charge density. While this approximation may not be that good (it is often very good), almost always enough of the structure is defined to solve the rest through one of a number of different techniques.

We do not need to know the phases that accurately; an average phase error of 10° or even 20° is good enough. Without any other information, the number of possible phase permutations is prohibitively large, of the order of 8^N for N beams and sampling at 45° intervals. We have the following additional information:

(i) The true charge density is real and positive at all points.

(ii) As mentioned above, the true charge density is zero outside a relatively small region normal to the surface. As discussed in the literature,[18–21] this reduces the possible phase permutations to at most 2^N.

(iii) Statistical (probabilistic) relationships connect the phases of different reflections.

The approach is to consider the possible values of the charge density that satisfy constraints (i) and (ii) above as sets, incorporating the constraints via mathematical projections. To this we add the set of all possible (complex) structures which have the experimental moduli. Finally, we consider the statistical relationships via both operators and functionals, and the set of fixed points (eigenvectors) or the sections of the functional below some value as a fourth set. The problem is then decomposed into finding the union of all the above sets; see, for instance, Fig. 1.

To start, instead of dealing with the true structure factors

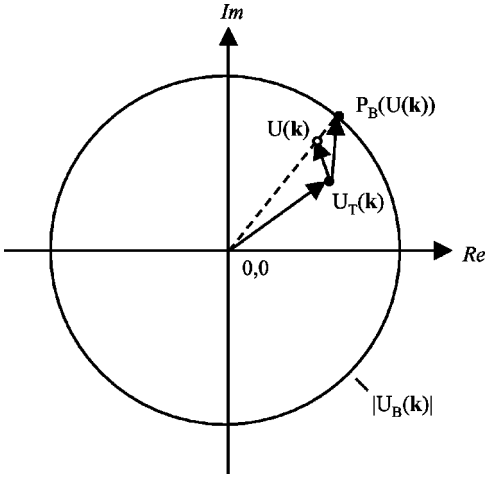

FIG. 2. For a given value of the bulk truncation rod $U_T(\mathbf{k})$ and some correct value $U(\mathbf{k})$ for the surface reflection, projection onto the experimental moduli $|U_B(\mathbf{k})|$ gives the vector from the end of $U_T(\mathbf{k})$ to $P_B[U(\mathbf{k})]$.

it is more convenient to use unitary structure factors defined by

$$U(\mathbf{k})=F(\mathbf{k})/\langle f(\mathbf{k})^2\rangle^{1/2}, \tag{1}$$

where the term on the right is the expectation value of the structure factor for random positions of atoms in the unit cell.

There exists a set of $U(\mathbf{k})$ values, S_M, all of which have the known moduli that are measured experimentally, i.e.,

$$|U(\mathbf{k})|=|U_e(\mathbf{k})| \tag{2}$$

the latter being the experimental values. A set is convex if any member lying on the line between two other members is also a member, i.e., if X and Y are two members of the set, it is convex if $\lambda X+(1-\lambda)Y$ for $0<\lambda<1$ is also a member; the above set S_M is not convex. From any value of $U(\mathbf{k})$ we can convert to values which lie on the set S_M using the projection operator

$$P_M[U(\mathbf{k})]=U(\mathbf{k})|U_e(\mathbf{k})|/|U(\mathbf{k})|, \tag{3}$$

i.e., replace the moduli by the measured values, leaving the phases unchanged.

For the second set, the charge density $u(\mathbf{r})$ [Fourier transform of $U(\mathbf{k})$] should be positive (a positivity constraint), and the set of all positive $u(\mathbf{r})$ is a convex set S_P. We can write this via the projection P_P where

$$\begin{aligned} P_P[u(\mathbf{r})]&=u(\mathbf{r}) \quad u(\mathbf{r})>0 \\ &=(1-\lambda)u(\mathbf{r}) \quad u(\mathbf{r})<0, \end{aligned} \tag{4}$$

where λ is called a relaxation parameter which lies in the range $0<\lambda<2$.[23–27] Values less than 1 are called under-relaxation, value greater than 1 over-relaxation. In general, values greater than 1 improve the convergence.

For the particular case of a three-dimensional surface, as mentioned in Sec. I the data are typically oversampled with an effective cell size of 40–60 Å normal to the surface. Within this cell there is only a limited region of 10–20 Å where the atomic positions are substantially different from

the bulk. This prior knowledge called (in the image processing literature) a support constraint, leads to a third, convex set S_S and a projection operator P_S where

$$P_S[u(\mathbf{r})]=u(\mathbf{r}), \quad -L<z<L$$
$$=0 \quad \text{otherwise}, \tag{5}$$

where $2L$ is the total width normal to the surface where relaxations have taken place. (We have experimented with incorporating a relaxation parameter here, but it does not appear to be particularly useful.) This is a particularly powerful constraint, with well-understood properties; the number of possible solutions when a one-dimensional support constraint is present is drastically reduced as mentioned earlier.

In many cases we have available measurements of the bulk rel-rods $|U_B(\mathbf{k})|$ for the reconstructed surface, and can make a good estimate of the structure factors (amplitude and phase) for a simple truncation of the surface, $U_T(\mathbf{k})$. Since we known that for the component of $U(\mathbf{k})$ that coincides with the bulk data

$$|U(\mathbf{k})+U_T(\mathbf{k})|=|U_B(\mathbf{k})|, \tag{6}$$

we can construct a projection of $U(\mathbf{k})$ (referred to as P_B) such that it will satisfy Eq. (6) (see Fig. 2) given by

$$P_B(U(\mathbf{k}))=\{U(\mathbf{k})+U_T(\mathbf{k})\}|U_B(\mathbf{k})|/|U(\mathbf{k})+U_T(\mathbf{k})|-U_T(\mathbf{k}). \tag{7}$$

The set of $U(\mathbf{k})$ values S_B that satisfies Eq. (6) is not convex.

Since the scattering of the x-rays comes from atoms, there are certain statistical phase relationships that can be established between different (typically the larger) $U(\mathbf{k})$'s. It is more convenient to write these in real space. The most probable vales of $u(\mathbf{r})$ are eigenfunctions (fixed points) of an operator T that represents these statistical relationships, i.e., for the true solution

$$T[u(\mathbf{r})]=u(\mathbf{r}) \tag{8}$$

and the difference between $u(\mathbf{r})$ and $T[u(\mathbf{r})]$ can be interpreted as a log likelihood that $u(\mathbf{r})$ comes from a set of atoms. An important property of T is whether it is contractive, i.e. for two different possible real space charge densities $x(\mathbf{r})$ and $y(\mathbf{r})$

$$\sum |T[x(\mathbf{r})]-T[y(\mathbf{r})]|^2 \leqslant \beta \sum |x(\mathbf{r})-y(\mathbf{r})|^2, \quad 0<\beta<1, \tag{9}$$

with the summation taken over r, or nonexpansive:

$$\sum |T[x(\mathbf{r})]-T[y(\mathbf{r})]|^2 \leqslant \sum |x(\mathbf{r})-y(\mathbf{r})|^2. \tag{10}$$

An important case of this is if we take

$$y(\mathbf{r})=T[x(\mathbf{r})], \tag{11}$$

so that

$$\sum |T[x(\mathbf{r})]-T[T[x(\mathbf{r})]]|^2 \leqslant \sum |x(\mathbf{r})-T[x(\mathbf{r})]|^2, \tag{12}$$

which implies that $T[x(\mathbf{r})]$ is closer to the solution than $x(\mathbf{r})$. Let us define this log-likelihood as equivalent to some functional $g(u)$, a figure of merit (FOM), which ideally should be zero. The sections of this functional such that

$$g(u)<\beta, \tag{13}$$

with β a constant, define a set S_T. If the operator is contractive this set, and the set of fixed points of T are convex; if the operator is nonexpansive both are in general nonconvex sets.

The mathematical problem to be solved now decomposes into finding the union of all the above sets, i.e., the ''feasible set'' S where

$$S=S_M \cap S_P \cap S_S \cap S_B \cap S_T. \tag{14}$$

If all the sets were convex, the feasible set is compact, and there is only one solution (plus, perhaps, a few minor variants). If one or more of the sets is nonconvex, the feasible set may be discontinuous, and we may have more than one feasible solution as initial models for the surface; see, for instance, Fig. 3. (Since S_M is never convex, we can never guarantee a single, unique solution.) The problem is known (e.g., Refs. 23 and 25) to be solvable by applying in some cyclic fashion the projections and the operator T; it is fully solvable to a unique solution if all the sets are convex, solvable in a locally convergent sense if they are not. For the latter case, one has to use a global search algorithm with different starting points.

For the operator, we need something which is correct for atoms. Starting with some estimate $u_n(\mathbf{r})$ for iteration cycle n, we use a normalized relative entropy[28,13,26] or Kullback-Leibler distance[29] as an operator:

$$u_{n+1}(\mathbf{r})=T_E(u_n(\mathbf{r}))=\alpha[u_n(\mathbf{r})\ln\{u_n(\mathbf{r})/\langle u_n(\mathbf{r})\rangle\}+\langle u_n(\mathbf{r})\rangle], \quad u_n(\mathbf{r})>0$$
$$=\langle u_n(\mathbf{r})\rangle, \quad u_n(\mathbf{r})<0, \tag{15}$$

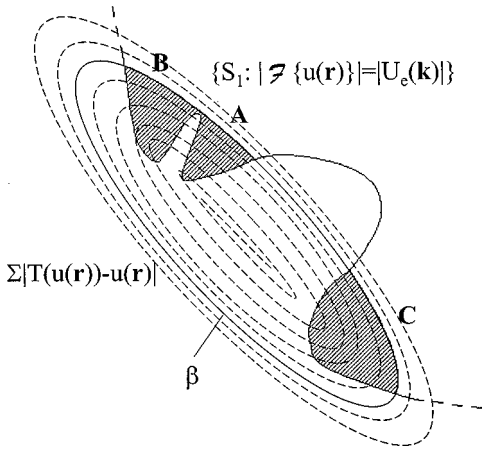

FIG. 3. Schematic illustration of the nonconvexity and discontinuous nature of the problem. Shown is the intersection for the nonconvex set and the set defined by $g(u) \leq \beta$ (assumed to be locally convex). The three subsets of the feasible set have been labeled *A*, *B*, and *C* since three such regions were found in some of the model systems, discussed later.

where $\langle u_n(\mathbf{r}) \rangle$ is the mean value, and α is a renormalization term defined as the value which minimizes

$$g(u) = \sum{}' \, |U_n(\mathbf{k}) - \alpha U_{n+1}(\mathbf{k})| \Big/ \sum{}' \, |U_n(\mathbf{k})|, \tag{16}$$

where Σ' indicates that we are not including the $k=0$ term and the summation is over the measured reflections only. Numerical tests demonstrate that the operator T_E is nonexpansive. For reference, Eqs. (8)–(12) earlier for the fixed points of the operator should only be interpreted for the measured reflections similar to Eq. (16). While atoms are fixed points of this operator for an infinite set of reflections, this is not the case for a finite set; typically measurements are only available for an ellipsoidal region with limited resolution perpendicular to the surface. To strengthen the operator, we introduce a window function $W(\mathbf{k})$,[13] and instead of Eq. (1) use the normalization:

$$U(\mathbf{k}) = W(\mathbf{k}) F(\mathbf{k}) / \langle f(\mathbf{k})^2 \rangle^{1/2}, \tag{17}$$

where $W(\mathbf{k})$ is real, positive, and a fixed point of T for an ellipsoidal region in three dimensions which includes, but is not limited to the measured reflection. (The fact that we include a complete ellipsoidal region here allows us to interpolate reflections, as will be discussed more below.) As a consequence, a combination of nonoverlapping atoms is a fixed point of T, a zero of $g(u)$ assuming all the atoms are the same. With measurement errors and different types of atoms we can only state that it is probable that the FOM is small at the correct solution.

Two other points need to be mentioned. In the above equations, some apply to the real-space charge density which implicitly contains all possible reflections; some, such as Eq. (16), apply only to the measured reflections. While only the experimentally determined phases are being used for the scaling term α and the FOM, the window function $W(\mathbf{k})$ is set up such that all the reflections within an ellipsoidal region are determined. Many of these are not measured, and so will be interpolated. Assuming that we have a reasonable number of measured reflections, we therefore have estimates of the phases of the unmeasured reflections. The set of all possible moduli for these is a convex set. Hence there is a very good probability that these will be relatively accurate interpolations, provided that they are correctly scaled. The same scaling term α used for the measured reflections should apply to the unmeasured ones, and in practice works very well. With all the experimental data that we have seen to date only a relatively small fraction of the possible beams have been measured. Without interpolation no such algorithm can be stable, but with both the positivity and (more importantly) the support constraint the algorithm is stable with a surprisingly small number of measurements.

Second, we know rather more than just the basic form of a statistical operator, we also know something about the statistical distribution. In particular, we know that the phase $\theta(\mathbf{k})$ for any reflection k in reciprocal space is

$$\theta(\mathbf{k}) \approx \theta(\mathbf{k}-\mathbf{h}) + \theta(\mathbf{h}), \tag{18}$$

with a probability distribution whose width scales inversely with $|U_e(\mathbf{k})|$, $|U_e(\mathbf{k}-\mathbf{h})|$, and $|U_e(\mathbf{h})|$. This implies that we can also use the predicted value of $|U(\mathbf{k})|$ as a gauge of the validity of a phase. We will refer to this as a ''phase-extension'' constraint, and code it into the algorithm as detailed below. The particular algorithm that we use is as follows.

(1) Start with some estimate (guess) for some of the phases, and couple this with the known moduli to give the initial estimate $U_0(\mathbf{k})$.

(2) For some given estimate of $u(\mathbf{r})$ (Fourier transforming as appropriate), i.e., $u_n(\mathbf{r})$ for cycle n, determine new estimates in reciprocal space in parallel using the statistical, positivity and support conditions via

$$U_n^E(\mathbf{k}) = F T_E[P_S(u_n(\mathbf{r})], \tag{19}$$

$$U_n^P(\mathbf{k}) = F P_P[P_S(u_n(\mathbf{r})], \tag{20}$$

where F stands for a Fourier transform.

(3) If we did not previously have an estimate for the phase for a particular k, apply the phase-extension constraint by setting

$$U_n^E(\mathbf{k}) = 0 \quad \text{if } |U_n^E(\mathbf{k})| < \gamma_n |U_e(\mathbf{k})|, \tag{21}$$

where γ_n is an adjustable scalar, and similarly for the positivity. Empirically, the form

$$\gamma_n = 0.3 \exp(-n/2) \tag{22}$$

is close to optimal.

(4) Combine the two estimates in parallel via

$$U_n^1(\mathbf{k}) = w U_n^E(\mathbf{k}) + (1-w) U_n^P(\mathbf{k}), \tag{23}$$

where a good choice of the weighting term w is

$$w = 0.5*[1 + \exp(-n/3)]. \tag{24}$$

(5) Correct the values of the moduli in reciprocal space back to the experimental values, using a relaxed projection operator

TABLE I. Unique atomic positions for model 1, $p2mm$ with $a=4.581$ Å, $b=18.325$ Å, and $c=64.79$ Å. For model 2, the y coordinate of the first atom is 0.128 726, leading to strong attenuation of the scattering from the first and last atoms, which lie almost exactly at bulk sites in the x-y plane. For model 3 an additional atom at (0, 0.25, 0.513) was added.

X	Y	Z
0.500 000	0.178 726	0.500 000
0.000 000	0.376 172	0.498 513
0.000 000	0.219 691	0.506 416
0.500 000	0.372 170	0.533 348

$$Q(\mathbf{k})=U_n(\mathbf{k})+\lambda[U_n^1(\mathbf{k})-U_n(\mathbf{k})], \tag{25}$$

$$\begin{aligned} U_{n+1}(\mathbf{k}) &= Q(\mathbf{k})|Ue(\mathbf{k})|/|Q(\mathbf{k})|, \quad \mathbf{k}\cap S_M \\ &= \alpha U_n^1(\mathbf{k}), \quad \mathbf{k}\cap S_U \\ &= 0 \quad \text{otherwise}, \end{aligned} \tag{26}$$

where S_M is the set of measured reflections and S_U the set of unmeasured reflections which lie within the aperture defined by the window function $W(\mathbf{k})$ of Eq. (17) and α the renormalization scaling from Eq. (16).

(6) Project (if appropriate) onto the bulk reflections.

(7) Evaluate via the FOM applied only to the relative entropy how well the current set of phases obeys both the statistical relationships, and how close are the moduli to the experimental values.

(8) If the FOM is decreasing, go back to (2) and continue the iteration; if it is increasing store the initial starting phases and best FOM.

The global search is then performed by a genetic algorithm[11] which finds approximate values for the phases in the feasible set. For reference, not all phases need to be specified in step (1), but typically only 5–10% of the stronger reflections.

III. BABINET SOLUTIONS

There is one interesting extension to the above case, which we will refer to as ''Babinet solutions.'' [12] To understand these, we will use the simpler unitary Sayre equation:

$$u(\mathbf{r})=Nu(\mathbf{r})^2, \tag{27}$$

which is true for N identical, nonoverlapping atoms. Let us decompose the charge density into two components; $u_M(\mathbf{r})$ for that due to the measured reflections (S_M), and $u_U(\mathbf{r})$ for that due to the unmeasured ones (S_U). Then

TABLE II. Atomic positions for model 4, cm with $a=20.79$ Å, $b=4.00$ Å, and $c=65.33$ Å, and $\alpha=\beta=\gamma=90.0$. The same unit-cell parameters were used for the fourth data set.

X	Y	Z
0.500 000	0.0	0.5
0.055 556	0.5	0.5
0.166 667	0.5	0.5
0.222 222	0.0	0.5

TABLE III. Rod data used for models 1 and 2. Shown are the largest l values ($l_{\max}$) in the data for given h and k values.

h	k	$l_{\max}$
0	2	20
0	3	25
0	5	30
0	6	30
0	7	30
0	9	30
0	11	30
0	13	30
0	14	30
1	2	27
1	3	30
1	5	32
1	6	32
1	7	32
1	9	30
1	10	30
2	1	30
2	2	30
2	3	30
2	5	30
2	6	29
2	7	30

$$u_M(\mathbf{r})=N(u_M(\mathbf{r})+u_U(\mathbf{r}))^2-u_U(\mathbf{r}). \tag{28}$$

Solutions of Eq. (28) are

(1) if $u_U(\mathbf{r})$ is small,

$$u_M(\mathbf{r})\approx 0 \quad \text{or} \quad 1/N; \tag{29}$$

(2) if $u_U(\mathbf{r})$ is large, approximately $1/N$,

$$u_M(\mathbf{r})\approx 0 \quad \text{or} \quad -1/N. \tag{30}$$

We start the iteration cycle with zero for the unmeasured reflections, so the algorithm will tend to find the solutions for case (1) above; the true solution might correspond to case (2). To handle this problem, we note that the primary difference between the two is reversal of the sign of $u_M(\mathbf{r})$—we are obtaining negative or Babinet solutions.

It might be effective to start with random initial values for some of the unmeasured reflections; we have not tested this. Instead, one can simply run a calculation starting with no

TABLE IV. Two-dimensional ($l=0$) data values used, except for when $k=4n$ with n an integer.

h	$k_{\max}$
0	19
1	19
2	18
3	15
4	11
5	3

TABLE V. Rod data used for models 3 and 4.

h	k	l_{max}
1	1	35
2	0	23
2	2	37
4	0	32
5	1	40
7	1	36
8	0	36
10	0	36

unmeasured moduli, then invert the phase of the measured values and rerun, retaining the estimates of the unmeasured reflections.

IV. NUMERICAL TESTS

Here we will present results for five different data sets, four model structures, and one set of experimental data. (We have also applied the method to three other sets of experimental data. However, since the structure of the latter have not previously been solved they will not be discussed further herein.) In addition to origin-defining reflections, in all cases a total of 80 reflections were quadrant permuted (90° steps) and the charge density was limited to a total height of 0.3 along the c axis. (Smaller heights will converge faster, but a conservative approach needs to be taken with real experimental data.) Within the genetic search algorithm[11] the population size was twice the number of bits, and the number of children twice this number. Each calculation was performed for ten generations, a total of about 5000 different cases, except for the first model where two different initial (random) seedings were used to improve the statistics.

TABLE VI. Two-dimensional ($l=0$) data values used for the last two data sets, excluding $h=3n$.

h_{max}	k
20	0
19	1
16	2
13	3

The four model structures contain atoms at the positions listed in Tables I and II, in standard crystallographic notation. The first two models are based around the 4×1 subcell of the InSb(001)2×8 reconstruction.[30] Comparing model 2 to model 1, the first and last atoms lie almost exactly at bulk sites in the x-y plane, with a height difference of about 2 Å. They will therefore only contribute strongly to relatively large l values, which are not well represented in the available reflections. This makes the second model harder to solve based just on the surface reflections, without using the bulk reflections. The third model includes an extra pair of atoms very close to two existing atoms. While physically unrealistic for true atomic positions, this structure no longer consists of well-separated atoms which therefore weakens the statistical relationships [the functional $g(u)$ discussed above will be small, but not zero for the correct solution]. The reflections used for the first three models are shown in Tables III and IV, again in standard crystallographic notation, and resemble those actually measured for the 4×1 subcell of the

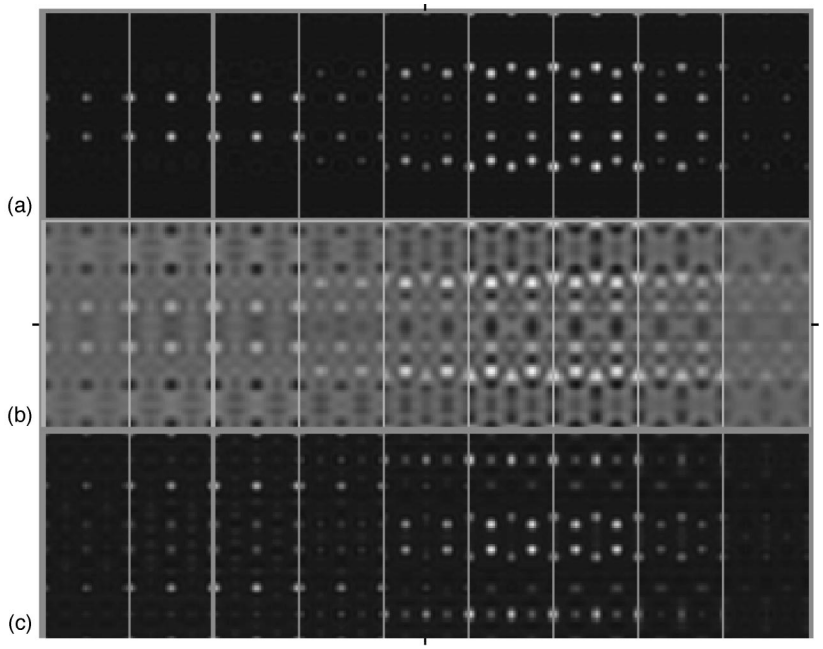

FIG. 4. Reconstructed charge densities sectioned normal to the z axis with slices of thickness approximately 0.5 Å, with white corresponding to regions of high charge density (atoms). In (a) and (c) the background level is close to zero, and in (b) negative holes (black regions) exist. Shown in (a) is the result with both the measured and unmeasured reflections, and in (b) the result with only the measured reflections. In 6(c) we show the best solution from set *C*. Strong artifacts in (b), some of which appear as ''atoms'' in (c), should be noted. For both (a) and (c) the sections were essentially featureless outside the regions shown; for (b) some artifacts persisted to substantial distances away from the atoms.

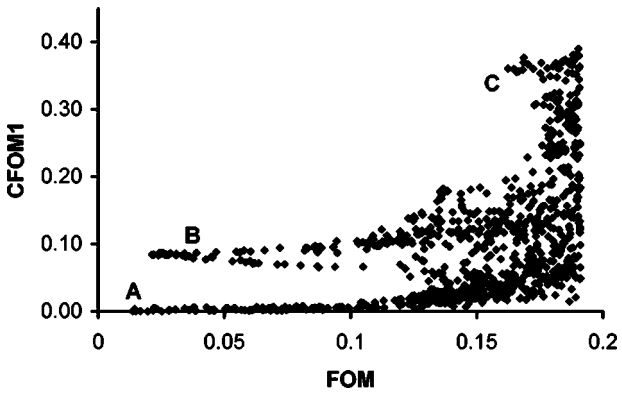

FIG. 5. Scatter plot of CFOM1 vs FOM for the top 1200 solutions for model 1. Three ''arrow'' shaped regions pointing towards smaller FOM values are marked as *A*, *B*, and *C* and correspond to three subsets of the total feasible set regime.

InSb(001)2×8 reconstruction.[30] The last two data sets are based on the Rb on Ge(111)3×1,[31] with the available reflections shown in Tables V and VI. Model 4 is a simple one-layer representation of the recently solved structure,[31] and the final data set is actual experimental data. To analyze in more detail how the algorithm performs, it is useful to introduce three parametrizations of the agreement between the reconstructed amplitudes and phases. The first is a consistency figure of merit (CFOM1) defined as

$$\mathrm{CFOM}=\sum U_t(\mathbf{k})\{1-\cos[\theta_t(\mathbf{k})-\theta_c(\mathbf{k})]\} \Big/ 2\sum U_t(\mathbf{k}), \tag{31}$$

where $\theta_t(\mathbf{k})$ are the true phases and $\theta_c(\mathbf{k})$ those returned by the algorithm, $U_t(\mathbf{k})$ is the true unitary structure factors, and the summation is taken over the reflections initially set via the genetic algorithm. The second, CFOM2, has the same equation but the sum is now extended to all reflections (except $\mathbf{k}=\mathbf{0}$). The last RFOM is defined as

$$\mathrm{RFOM}=\sum |U_t(\mathbf{k})-U_c(\mathbf{k})| \Big/ \sum U_t(\mathbf{k}). \tag{32}$$

The first metric CFOM1 will describe how good a match is obtained between the phases of the reflections used in the starting set and their true values, being zero for perfect agreement. Based upon prior experience, a value of approximately 0.1 or better will give a very good reconstruction of the charge density. The second metric does the same for all the phases, i.e., it includes the phases of the unmeasured reflections as well. Since phases are more important than the moduli in obtaining a viable reconstruction, values of 0.1 or better are very good. The final metric RFOM shows how good an overall fit is achieved including interpolation of the moduli of the unmeasured reflections. This need not be that small for a good restoration of the charge density.

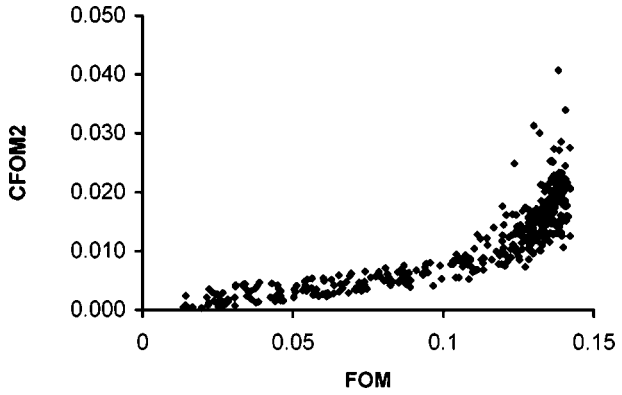

FIG. 6. Scatter plot of CFOM2 vs FOM for model 1, demonstrating that small FOM values give very good restoration of the phases of the unmeasured reflections.

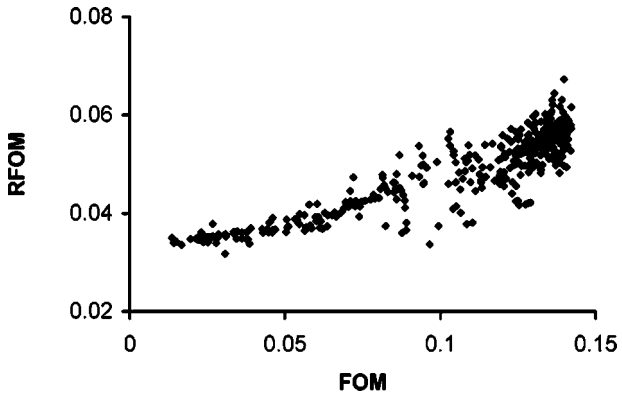

FIG. 7. Scatter plot of RFOM vs FOM for model 1, showing that a reasonably good restoration of the amplitudes is obtained.

The first model structure solves without use of any bulk projections, and a typical result is shown in Fig. 4(a). Not only are the atom sites very well resolved, very good reconstruction of the unmeasured reflections is obtained. To illustrate this, Fig. 4(b) shows the charge density if only the measured reflections are used. Figure 5 plots CFOM1 versus the FOM for the best 1200 solutions for a search of about 10 000. As can be seen from the figure, there are three arrowlike features in the scatter plot pointing towards low FOM values, indicating three different subsets (solutions) within the total feasible set. The first two, labeled *A* and *B* in the figure, both give very good reconstructions similar to Fig. 4(a). The third, labeled *C*, is rather inferior and is shown in Fig. 4(c). At a rigorous level it would be necessary to check (via a χ^2 refinement) each of these three possible solutions for a full structure determination. Shown in Fig. 6 is CFOM2 versus FOM, and shown in Fig. 7 is RFOM versus FOM for the top 120 solutions. Good results are obtained for restoration of the phases of the unmeasured reflections, not so good (but adequate) for restoration of the moduli as well. To show the later in more detail, Fig. 8 is a plot of the true versus restored moduli for one of the stronger rel-rods (i.e., *h* and *k* fixed, and different values of *l*).

The second model structure does not solve so well using the surface reflections, as can be seen from the plot of CFOM1 versus FOM in Fig. 9; the incorrect set of solutions marked *C* [similar to the subset *C* in Fig. 5) move to lower FOM values. This is accentuated in the third model, where subset *C* moves to lowest FOM as shown in Fig. 10(a).] While part of the structure is correct (the nearly coincident atoms are found), the second layer is too close to a perfect bulk and is not determined. Sufficient information might be available for the true structure to be determined by subsequent structure completion by Fourier methods, χ^2 analysis, and chemical information, but this would not be easy. How-

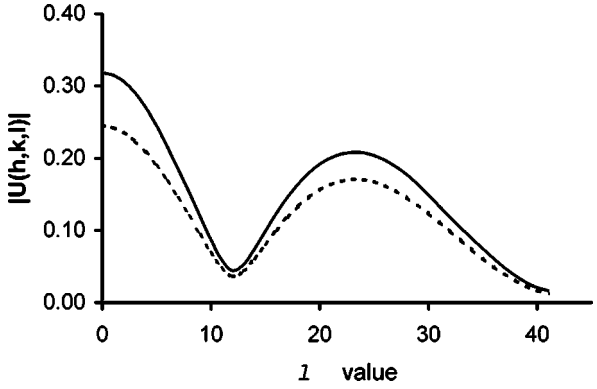

FIG. 8. Plot of the $(0,4,l)$ rel-rod showing both the true values (solid line) and the restored values.

ever, when three bulk rods [the $(1,0,l)$, $(0,4,l)$, and $(1,4,l)$ rods] are included in the calculation, it does solve as shown by the plot of CFOM1 versus FOM in Fig. 10(b). (The second model behaves similarly, collapsing to the correct solution when projection onto the bulk rods is employed.)

As an alternative approach, the complexity of the problem can be reduced by solving first in two dimensions (both for the phases and a χ^2 refinement) prior to analyzing the three-dimensional (3D) problem. One can then apply some of the phases from the 2D solution as part of the starting set. For both models 2 and 3 this works well, and reduces substantially the FOM of the correct solution within the feasible set as illustrated in Fig. 11.

The final two data sets demonstrate the importance of Babinet solutions. The structure is equivalent to one Ge bilayer (at the same height) minus a dimer. The standard algorithm suggests a dimer structure as shown in Fig. 12(a) for model 4, while the Babinet solution in Fig. 12(b) gives the correct solution. The same is found with the experimental data [Fig. 12(c)], very close to that recently determined,[31] although see Sec. V. (In practice, both the dimer and four-atom solutions would have to be considered as part of the feasible set of possible solutions for a subsequent full refinement.) Note that if the bulk rods were included in the analysis, the possibility of a Babinet solution would be removed.

V. DISCUSSION

It is perhaps best to start with a word of warning: no method is completely foolproof, and it is always possible to

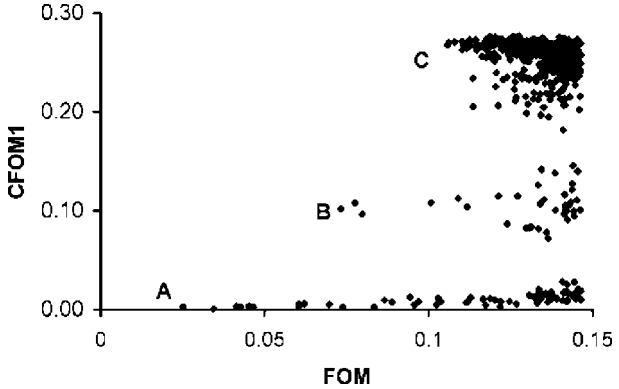

FIG. 9. Scatter plot of CFOM1 vs FOM for model 2; the incorrect subset of solutions marked *C* have moved to lower FOM values relative to Fig. 5.

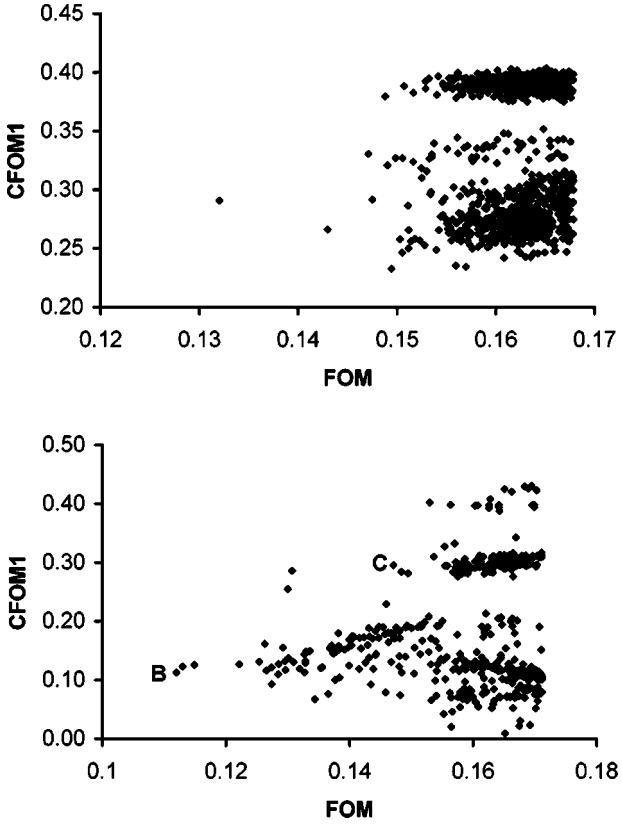

FIG. 10. Scatter plots of CFOM1 vs FOM for model 3. In (a) only the surface data are used, and the incorrect solution *C* found in both Figs. 5 and 9 dominates. However, when projection onto the bulk rods is exploited only the true move to smaller FOM values as shown in (b).

end up with the wrong result if due care is not exercised. While the algorithms herein are powerful, care is needed with the global search and in some cases there may simply not be enough data to determine a surface structure. (If there are not enough data for direct methods to work, there are almost certainly not enough data to solve the full three-dimensional structure; one or two measurements of the rel-rods may not be adequate.)

Very important is the quality of the data, both in terms of

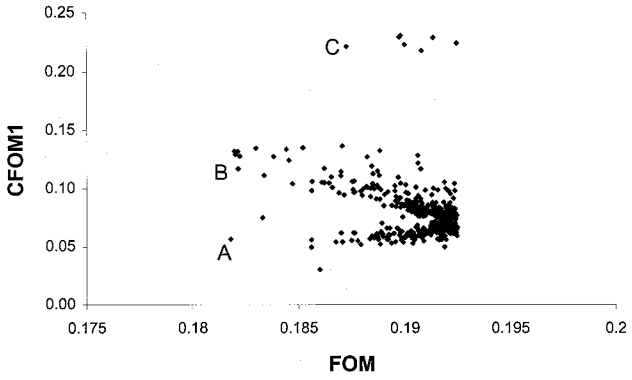

FIG. 11. Scatter plot of CFOM1 vs FOM for model 3, where ten $(h,k,0)$ reflections have been fixed to mimic initial 2D solutions. Solutions *A* and *B* dominate, albeit with higher FOM values than in Fig. 10(a).

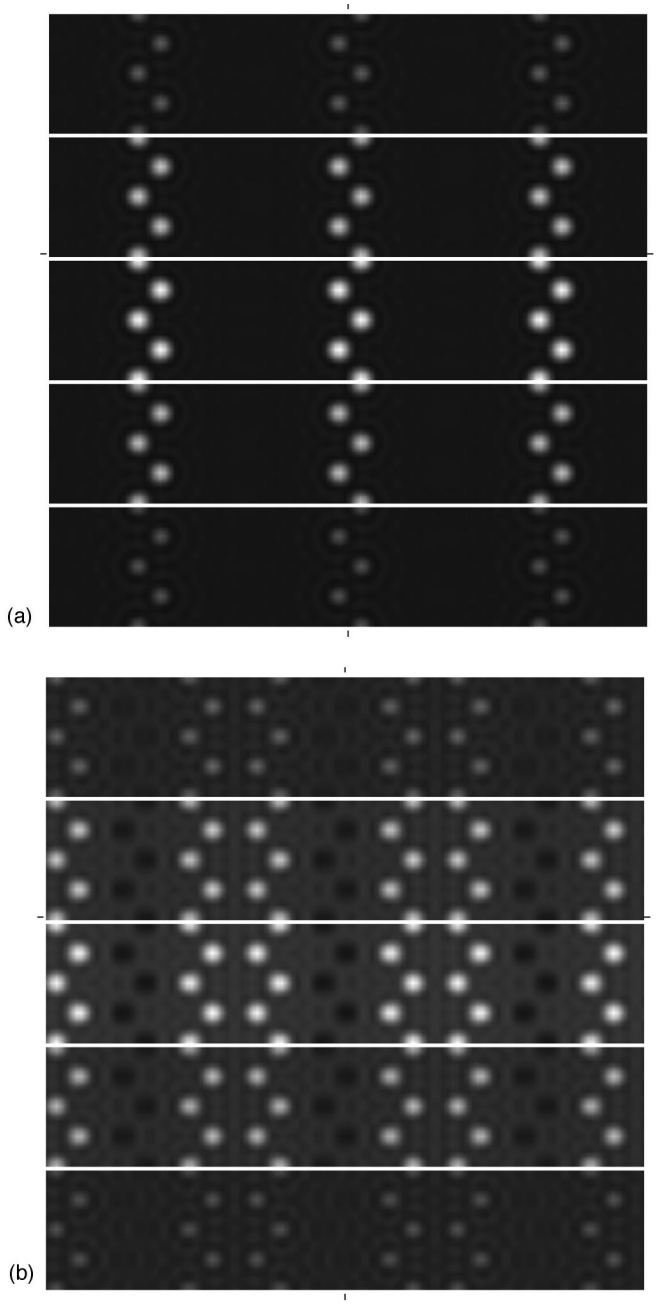

FIG. 12. Charge-density sections normal to the z axis with a slice thickness of about 0.5 Å for model 4, and the experimental data. Shown in (a) is a dimer solution obtained with the standard algorithm; in (b) we show the Babinet solution after a sign reversal projection as described in the text, and in (c) the Babinet solution for the experimental data.

measurement errors and other issues such as secondary phases. It is not at all unusual to have defects and partial occupancies of certain sites due to surface disorder. Furthermore, in almost all cases there will also be additional surface phases whose reflections may be coincident with that of the target surface structure. The latter means that it may be difficult to extract the bulk component due to a particular surface structure (or domain), although in principle it is possible to extend the method herein to include simultaneous projections onto a number of different surface structures.

Also of prime importance is the symmetry of the reconstruction. This severely constrains the possible solutions, and might be determined by STM, although rather little attention has been paid to this to date. Better information about the unit cell contents always helps, since this is used in the normalizations. (In practice one can run for a number of different assumptions about the symmetry and contents.)

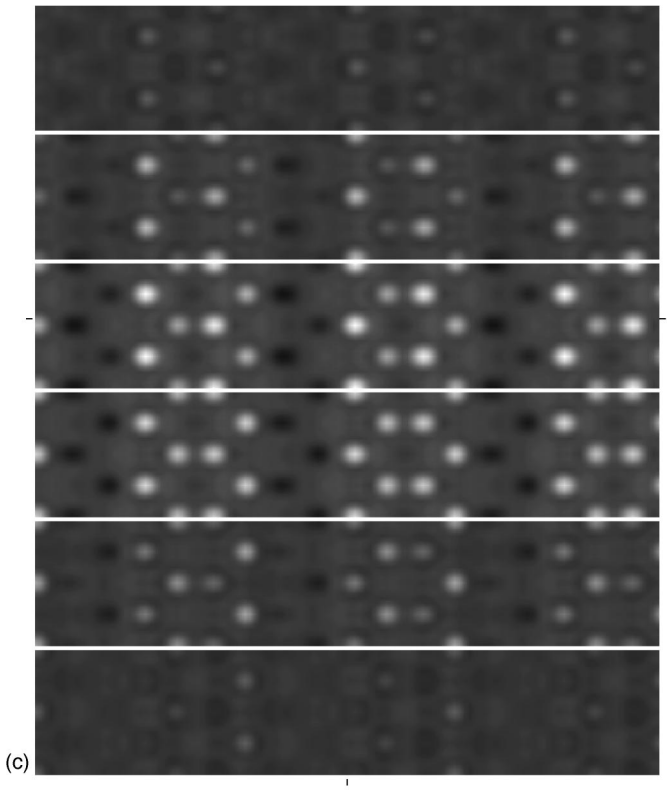

FIG. 12. (*Continued*).

Many other versions of direct methods will probably also work for surfaces in three dimensions, and it is not unreasonable to hope for fully automated structure solutions in a few years. Some modification to existing codes will almost certainly be needed, particularly the introduction of the support constraint. At present some manual intervention is still required in an interpretation of the maps and a conversion from these to atomic positions, a problem since bias can be introduced. Due to the large number of unmeasured reflections conventional Fourier difference techniques do not work very well, although some preliminary results suggest that modifications to include projections may be viable.

These caveats aside, we have outlined herein what appears to be a completely general method of solving, without guesswork, surface structures using three-dimensional x-ray-diffraction data. In many cases the 3D data will solve on its own (with appropriate attention paid to Babinet solutions). As an alternative, if bulk-rod data are available and the coverage of the structure relatively is well known (for scaling purposes), this information can be used; if a reasonable χ^2 can be obtained in two dimension, a determination of the full structure from a limited number of rods is quite realistic.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation Grant No. DMR-9705081.

[1] G. Giacovazzo, *Direct Methods in Crystallography* (Plenum, New York, 1980).
[2] G. Bricogne, Acta Crystallogr., Sect. A: Found. Crystallogr. **40**, 410 (1984).
[3] M. M. Woolfson, Acta Crystallogr., Sect. A: Found. Crystallogr. **43**, 593 (1987).
[4] G. M. Sheldrick, Acta Crystallogr., Sect. A: Found. Crystallogr. **46**, 467 (1990).
[5] M. M. Woolfson and H. Fan, *Physical and Non-Physical Methods of Solving Crystal Structures* (Cambridge University Press, Cambridge, 1995).
[6] D. L. Dorset, Acta Crystallogr., Sect. A: Found. Crystallogr. **52**, 753 (1996).
[7] C. J. Gilmore, Acta Crystallogr., Sect. A: Found. Crystallogr. **52**, 561 (1996).
[8] L. D. Marks, R. Plass, and D. L. Dorset, Surf. Rev. Lett. **4**, 1 (1997).
[9] C. J. Gilmore, L. D. Marks, D. Grozea, C. Collazo-Davila, E. Landree, and R. Twesten, Surf. Sci. **381**, 77 (1997).
[10] C. Collazo-Davila, L. D. Marks, K. Nishii, and Y. Tanishiro, Surf. Rev. Lett. **4**, 65 (1997).
[11] E. Landree, C. Collazo-Davila, and L. D. Marks, Acta Crystallogr., Sect. B: Struct. Sci. **53**, 916 (1997).
[12] L. D. Marks, E. Bengu, C. Collazo-Davila, D. Grozea, E. Landree, C. Leslie, and W. Sinkler, Surf. Rev. Lett. **5**, 1087 (1998).
[13] L. D. Marks and E. Landree, Acta Crystallogr., Sect. A: Found. Crystallogr. **54**, 296 (1998).
[14] C. Collazo-Davila, D. Grozea, and L. D. Marks, Phys. Rev. Lett. **80**, 1678 (1998).
[15] D. Grozea, E. Landree, L. D. Marks, R. Feidenhans'l, M. Nielsen, and R. L. Johnson, Surf. Sci. **418**, 32 (1998).
[16] C. Collazo-Davila, D. Grozea, L. D. Marks, R. Feidenhans'l, M. Nielsen, L. Seehofer, L. Lottermoser, G. Falkenberg, and R. L. Johnson, Surf. Sci. **418**, 395 (1998).
[17] Richard Plass, Kenneth Egan, and Marija Gajdardziska-Josifovska, Chris Collazo-Davila, Daniel Grozea, Eric Landree, and Laurence D. Marks, Phys. Rev. Lett. **81**, 4891 (1998).
[18] J. R. Fienup, Opt. Lett. **3**, 27 (1978).
[19] J. R. Fienup, J. Opt. Soc. Am. A **4**, 118 (1987).
[20] J. C. Dainty and J. R. Fienup, in *Image Recovery: Theory and Application*, edited by H. Stark (Academic, Orlando, 1987).
[21] M. H. Hayes, in *Image Recovery, Theory and Application* (Ref. 20).
[22] L. G. Gubin, B. T. Polyak, and E. V. Raik, USSR Comput. Math. Math. Phys. **7**, 1 (1967).
[23] D. C. Youla, in *Image Recovery, Theory and Application*, edited by H. Stark (Academic, Orlando, 1987).
[24] M. I. Sezan, Ultramicroscopy **40**, 55 (1992).
[25] P. L. Combettes, Adv. Imaging Electron Phys. **95**, 155 (1996).
[26] L. D. Marks, W. Sinkler, and E. Landree, Acta Crystallogr., Sect. A: Found. Crystallogr. (to be published).
[27] A. Levi and H. Stark, in *Image Recovery, Theory and Application* (Ref. 23).
[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[29] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[30] R. Feidenhans'l (private communication).
[31] L. Lottermoser, E. Landemark, D.-M. Smilgies, N. Nielsen, R. Feidenhans'l, G. Falkenberg, R. L. Johnson, M. Gierer, A. P. Seitsonen, H. Kleine, H. Bludau, H. Over, S. K. Kim, and F. Jona, Phys. Rev. Lett. **80**, 3980 (1998).